AFRL-DE-PS-JA-2007-1011

AFRL-DE-PS-JA-2007-1011

Biased Cramer-Rao lower bound calculations for inequality-constrained estimators (Preprint)

Charles Matson
Alim Haji

1 September 2006

Journal Article

AIR FORCE RESEARCH LABORATORY
Directed Energy Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
1 September 06	Journal Article	Oct 04-Aug 06

4. TITLE AND SUBTITLE Biased Cramer-Rao lower bound calculations for inequality-constrained estimators (Preprint)	**5a. CONTRACT NUMBER** F29601-01-D-0083 TO 06 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER** 061102F
6. AUTHOR(S) Charles L. Matson, Alim Haji	**5d. PROJECT NUMBER** D0D0 **5e. TASK NUMBER** B3 **5f. WORK UNIT NUMBER** AD
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AFRL/RDSA 3550 Aberdeen Avenue SE Kirtland AFB, NM 87117-5776	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory 3550 Aberdeen Ave. SE Kirtland AFB, NM 87117-5776	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/RDSA **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AFRL-DE-PS-JA-2007-1011

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited.

13. SUPPLEMENTARY NOTES

Submitted for publication in JOSA A, Vol. 23, No. 11, November 2006.

GOVERNMENT PURPOSE RIGHTS

14. ABSTRACT

Unbiased Cramer-Rao lower bound (CRB) theory can be used to calculate lower bounds to the variances of unbiased estimates of a set of parameters given only the probability density function of a random vector conditioned on the true parameter values. However, when the estimated parameter values are required to satisfy inequality constraints such as positivity, the resulting estimator is typically biased. To calculate CRBs for biased estimates of the parameter values, an expression for the bias gradient matrix must also be known. Unfortunately, this expression often does not exist. Because expressions for biased CRBs are preferable to sample variances calculations, alternate methods for deriving biased CRB expressions associated with inequality constraints are needed. Here we present an alternate approach that is biased upon creating the probability density function associated with a given biased estimate of these parameters using the available knowledge of estimator properties. We apply this approach to the calculations of biased CRBs for estimators that use a positivity constraint for a specific measurement model and discuss the benefits and limitations of this approach.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Charles Matson
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	SAR	12	**19b. TELEPHONE NUMBER** *(include area code)*

Biased Cramér-Rao lower bound calculations for inequality-constrained estimators

Charles L. Matson and Alim Haji
Air Force Research Laboratory, Kirtland AFB, NM 87117-5776

Abstract

Unbiased Cramér-Rao lower bound (CRB) theory can be used to calculate lower bounds to the variances of unbiased estimates of a set of parameters given only the probability density function of a random vector conditioned on the true parameter values. However, when the estimated parameter values are required to satisfy inequality constraints such as positivity, the resulting estimator is typically biased. To calculate CRBs for biased estimates of the parameter values, an expression for the bias gradient matrix must also be known. Unfortunately, this expression often does not exist. Because expressions for biased CRBs are preferable to sample variance calculations, alternate methods for deriving biased CRB expressions associated with inequality constraints are needed. Here we present an alternate approach that is based upon creating the probability density function associated with a given biased estimate of these parameters using the available knowledge of the estimator properties. We apply this approach to the calculation of biased CRBs for estimators that use a positivity constraint with and without a support constraint for a specific measurement model and discuss the benefits and limitations of this approach.

1. Introduction

When seeking to estimate a vector of unknown parameters from a set of measurements, it can be useful to be able to calculate lower bounds to the variances of any estimates of these parameters. One potential use of these bounds is to determine, a priori, whether or not the desired estimation accuracy can be achieved for a given measurement model. Another use is to employ them in designing an experiment for which it is known that the desired accuracy can be achieved. A third use for these types of bounds is to check the accuracy of a software implementation of an estimation algorithm by comparing its results to the lower bounds.

One well-known theory, called the unbiased Cramér-Rao lower bound (CRB) theory,[1] provides algorithm-independent lower bounds (called unbiased CRBs) to the variances of any unbiased estimates of a vector of parameters obtained from a set of measurements. These bounds can be calculated using only the probability density function (PDF) of the measurements that is a function of the true parameter values. The unbiased CRBs may not be achievable, but often they can be achieved or at least approached closely. However, in many cases, unbiased parameter estimates are not possible (non-invertible system models) or desirable (excessive noise amplification in the unbiased estimates). For biased parameter estimates, biased CRBs can be calculated.[1] In addition to the measurement PDF, the bias gradient matrix of the parameter estimates is needed to calculate biased CRBs. We note that an alternate approach to biased CRB theory is available that generates biased CRBs (called uniform CRBs) that depend upon only the norm of the bias gradient and not its direction.[2]

Our interest in CRB theory is to use these lower bounds to quantify how the various types of knowledge included in parameter estimation algorithms affect the quality of the parameter estimates. There exists a large number of possible types of knowledge that can be included in the estimation process (c.f. Ref.[3]), depending upon the type of input signal and the system through which it passes prior to measurements being taken. Two types of knowledge that have been used extensively in parameter estimation algorithms are the support constraint (restricting the solution set to parameters that are only non-zero in a specified region) and the positivity constraint (also known as non-negativity; i.e., restricting the solution set to parameters that are not negative). The support constraint is a specific example of a more general class of constraints known as equality constraints, where functions of the noise-free parameters are known to be zero. Similarly, the positivity constraint is a specific example of a class of constraints known as inequality constraints, where functions of the noise-free parameters are known to be greater than zero.

In previously-published seminal work,[4] a detailed CRB theory was developed that quantified the ability of equality and inequality constraints to reduce noise in parameter estimates for estimators with a given bias gradient matrix. One key finding in this work is that imposing equality constraints on the parameters to be estimated can produce lower CRBs than without the equality constraints. A second key finding is that inequality constraints on the parameters to be estimated cannot produce lower CRBs as long as the bias gradient matrix remains unchanged. This second finding is at first perplexing, because it is well known to signal processing practitioners that a positivity constraint often produces less noisy signal reconstructions than possible without a positivity constraint. The answer to this conundrum is that the variance reduction brought about by inequality constraints on the parameters is

AFRL/DE 06-389

where ln denotes the natural logarithm and E[] denotes the expected value of the quantity in the brackets. Here we assume that the FIM is invertible; if not, either a subset of or functions of the parameters may be estimable in an unbiased way.[6] As can be seen from Eq.(1), the FIM is a function only of the measurement PDF and thus so are the unbiased CRBs.

Biased CRBs can be calculated in terms of the FIM defined in Eq.(1) using the following expression

$$\mathrm{CRB}_{\mathrm{biased}}(\theta) = \mathrm{diag}\left\{\left[\mathbf{I} + \nabla_\theta \mathbf{b}(\theta)\right]\mathbf{F}^{-1}(\theta)\left[\mathbf{I} + \nabla_\theta \mathbf{b}(\theta)\right]^T\right\}, \tag{2}$$

where $\mathbf{b}(\theta)$ is the bias and $\nabla_\theta \mathbf{b}(\theta)$ is the bias gradient matrix associated with a given estimator that is a function of θ, and $\mathbf{I}$ is the identity matrix of appropriate dimension. Obviously, an expression for $\nabla_\theta \mathbf{b}(\theta)$ is needed to carry out the calculations in Eq.(2) in addition to the measurement PDF.

When it is known that the true parameter values satisfy a set of equality constraints, the FIM can be modified to include this additional information and improve the accuracy of both unbiased and biased CRBs. The general theory on how to modify the FIM to include equality constraints can be found in Ref. [4]. For the special case of a support constraint, the constrained FIM is obtained by zeroing out the rows and columns of the unconstrained FIM corresponding to the elements of θ outside the known support. Then $\mathbf{F}^{-1}$ in Eq.(2) is replaced by its pseudoinverse. No changes need to be made to $\nabla_\theta \mathbf{b}(\theta)$.

When it is known that the true parameter values satisfy a set of inequality constraints, the FIM remains unchanged; however, the estimator bias gradients are altered when the inequality constraints are active. In this case, $\nabla_\theta \mathbf{b}(\theta)$ must be modified to have the biased CRBs reflect the inequality constraint. Unlike for equality constraints, no general theory is available that describes how to modify an existing $\nabla_\theta \mathbf{b}(\theta)$ to include the inequality constraints. Finally, when both equality and inequality constraints are jointly known about the true parameter values, the associated CRBs are calculated by separately modifying the unconstrained FIM to reflect the equality constraints and the bias gradient matrix to reflect the inequality constraints.

3. Estimation PDF creation

In this section, we describe first how estimation PDFs can be used to obtain CRB expressions that are equivalent to the standard CRB expressions that are obtained using measurement PDFs. Second, we show how to create estimation PDFs for three different estimators: (1) one that employs a positivity constraint, (2) one that employs a support constraint, and (3) one that employs both a positivity and a support constraint. For all three cases, we assume that the FIM is invertible. Third, we describe conceptually how to extend the estimation PDF creation method to more general inequality and equality constraints as well as to non-invertible FIMs. Finally, we use our method to create estimation PDFs for a simple measurement model and the three different estimators described in this section in order to demonstrate how to apply the estimation PDF creation process. In Section 4, these expressions will be used to calculate and evaluate CRBs.

3.1 CRB calculations using estimation PDFs

To motivate the benefits and validity of creating a estimation PDF for CRB calculations, consider first a standard result from probability theory.[7] Let $\mathbf{y}$ be a random vector and $\mathbf{g}(\mathbf{y})$ be an arbitrary vector function of $\mathbf{y}$. The expected value of $\mathbf{g}$ can be calculated in one of two equivalent ways:

$$\begin{aligned} E[\mathbf{g}(\mathbf{y})] &= \int \mathbf{g}(\mathbf{y}) f(\mathbf{y};\theta)\, d\mathbf{y} \\ &= \int \mathbf{z}\, f(\mathbf{z};\theta)\, d\mathbf{z} \end{aligned}, \tag{3}$$

where $\mathbf{z}=\mathbf{g}(\mathbf{y})$ and $f(\mathbf{z};\theta)$ is the PDF of $\mathbf{z}$ that is a function of θ. Because $\mathbf{z}$ is the independent set of variables in $\mathbf{z}$ space, no knowledge of $\mathbf{g}(\mathbf{y})$ is necessary for calculating the desired expected value when using the second integral expression in Eq.(3). In other words, $f(\mathbf{z};\theta)$ incorporates all the information contained in $\mathbf{g}(\mathbf{y})$. When a mathematical expression for $\mathbf{g}(\mathbf{y})$ is available, $f(\mathbf{z};\theta)$ can be calculated from $f(\mathbf{y};\theta)$ using standard PDF transformation methods.[7,8]

This same principle can be used to create a biased CRB expression that is equivalent to the expression given in Eq.(2). To use Eq.(2), an expression for the bias gradient matrix $\nabla_\theta \mathbf{b}(\theta)$ must be available. In practice,

3.2.1 Positivity-constrained estimation PDF creation

Let $\hat{\theta}_p$ be a positivity-constrained estimator of θ with PDF $f_p(\hat{\theta}_p;\theta)$. We define $\hat{\theta}_p$, element by element, as follows: we set the j^{th} element of $\hat{\theta}_p$, $\hat{\theta}_{pj}$, equal to the j^{th} element of $\hat{\theta}_u$, $\hat{\theta}_{uj}$, if $\hat{\theta}_{uj}$ is non-negative, otherwise we set $\hat{\theta}_{pj}$ equal to zero. Of course, this is not the only way that positivity can be enforced. Other ways include replacing negative numbers with their absolute values or adding a positive constant that is larger than the absolute value of the largest-magnitude negative number. The use of one of these alternate ways to enforce positivity will change the PDF of the positivity-constrained estimator and thus change the CRBs. This issue is not explored further here because the focus of this paper is on the PDF-creation process, not on the optimal way to enforce positivity. In addition, we require that the statistical properties of $\hat{\theta}_p$ be consistent with the statistical properties of $\hat{\theta}_u$ as much as possible while satisfying the positivity constraint. The latter requirement can be understood by realizing that a constrained estimate of a set of parameters is chosen to be as consistent as possible with the unconstrained estimate while also satisfying the known constraints. We will now create $f_p(\hat{\theta}_p;\theta)$ based upon this description of the estimator. When all the elements of $\hat{\theta}_u$ are non-negative, the positivity constraint is not active and thus $f_p(\hat{\theta}_p;\theta)$ = $f_u(\hat{\theta}_u;\theta)$ for these values. However, when one or more elements of $\hat{\theta}_u$ are negative, the positivity constraint is active and this equality does not hold. When the vector θ is N-dimensional, there are 2^N-1 possible combinations of the elements of $\hat{\theta}_u$ for which the positivity constraint is active. As a result, there are 2^N terms in the expression for $f_p(\hat{\theta}_p;\theta)$. We will first create an expression for the term in $f_p(\hat{\theta}_p;\theta)$ that corresponds to the case where only the first element of $\hat{\theta}_u$, $\hat{\theta}_{u1}$, is negative. Based upon this result, we will then present a general expression that can be used to generate any term in $f_p(\hat{\theta}_p;\theta)$ except for the case where all the elements of $\hat{\theta}_u$ are negative. This last case can be obtained by subtracting the integral of the sum of all the other terms of $f_p(\hat{\theta}_p;\theta)$ from one since PDFs integrate to one. We note that the number of terms in $f_p(\hat{\theta}_p;\theta)$ can become quite large. In Section 5, we address potential ways to deal with this problem.

The PDF creation process can be achieved with the help of the PDF form of Bayes' theorem.[7] When only the first element of $\hat{\theta}_p$ has been set equal to zero because of the positivity constraint, Bayes' theorem can be used to express the constrained joint PDF $f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)$ in terms of its associated constrained conditional PDF $f_p(\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN}|\hat{\theta}_{p1}=0;\theta)$ and constrained marginal PDF $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)\,d\hat{\theta}_{p2}\cdots d\hat{\theta}_{pN}$ as follows:[7]

$$f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)=f_p(\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN}|\hat{\theta}_{p1}=0;\theta)\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)\,d\hat{\theta}_{p2}\cdots d\hat{\theta}_{pN}\,. \quad (7)$$

Based upon Eq.(7), we can create $f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)$ by creating separately $f_p(\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN}|\hat{\theta}_{p1}=0;\theta)$ and $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_p(\hat{\theta}_{p1}=0,\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN};\theta)\,d\hat{\theta}_{p2}\cdots d\hat{\theta}_{pN}$ and then multiplying the two expressions together.

From a closer look at Eq.(7), we can see that the constrained conditional PDF carries the statistical information about $\hat{\theta}_p$ when $\hat{\theta}_{p1}=0$ while the constrained marginal PDF is the integral over all the values of $\hat{\theta}_p$ for which $\hat{\theta}_{p1}=0$. We desire to have the constrained joint PDF be as consistent with the unconstrained joint PDF as possible while satisfying the known constraints. Because the unconstrained conditional PDF describes the statistical properties of the unconstrained estimate of θ, we satisfy the unconstrained PDF consistency desire by setting $f_p(\hat{\theta}_{p2},\ldots,\hat{\theta}_{pN}|\hat{\theta}_{p1}=0;\theta)$ equal to the unconstrained conditional PDF $f_u(\hat{\theta}_{u2},\ldots,\hat{\theta}_{uN}|\hat{\theta}_{u1}=0;\theta)$. The difference between the constrained and unconstrained joint PDFs is thus carried in the marginal PDFs. We obtain the

$$
\begin{aligned}
f_p\left(\hat{\theta}_{p1}=0,\ldots,\hat{\theta}_{pM}=0,\hat{\theta}_{p(M+1)},\ldots,\hat{\theta}_{pN};\theta\right) &= f_u\left(\hat{\theta}_{u(M+1)},\ldots,\hat{\theta}_{uN}\,\middle|\,\hat{\theta}_{u1}=0,\ldots,\hat{\theta}_{uM}=0;\theta\right) \\
&\times\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_u\left(\hat{\theta}_{u1}=0,\ldots,\hat{\theta}_{uM}=0,\hat{\theta}_{u(M+1)},\ldots,\hat{\theta}_{uN};\theta\right)d\hat{\theta}_{u(M+1)}\cdots d\hat{\theta}_{uN}\right]^{-1} \\
&\times\int_{-\infty}^{0}\cdots\int_{-\infty}^{0}\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_u\left(\hat{\theta}_{u1},\ldots,\hat{\theta}_{uN};\theta\right)d\hat{\theta}_{u(M+1)}\cdots d\hat{\theta}_{uN}\right]d\hat{\theta}_{u1}\cdots d\hat{\theta}_{uM} \\
&\times\delta\left(\hat{\theta}_{p1}\right)\cdots\delta\left(\hat{\theta}_{pM}\right)u^+\left(\hat{\theta}_{p(M+1)}\right)\cdots u^+\left(\hat{\theta}_{pN}\right)
\end{aligned}
\quad . \quad (11)
$$

Because M is arbitrary and the elements of $\hat{\theta}_u$ can be arbitrarily reordered, Eq.(11) can be used to calculate all the terms that are part of $f_p\left(\hat{\theta}_p;\theta\right)$ except for the cases of no negative elements and all negative elements. The term corresponding to no negative elements in $\hat{\theta}_u$, $f_p\left(\hat{\theta}_{p1},\ldots,\hat{\theta}_{pN};\theta\right)$, is given by

$$
f_p\left(\hat{\theta}_{p1},\ldots,\hat{\theta}_{pN};\theta\right)=f_u\left(\hat{\theta}_{u1},\ldots,\hat{\theta}_{uN};\theta\right)u^+\left(\hat{\theta}_{p1}\right)\cdots u^+\left(\hat{\theta}_{pN}\right), \quad (12)
$$

and the term corresponding to all negative elements $\hat{\theta}_u$, $f_p\left(\hat{\theta}_{p1}=0,\ldots,\hat{\theta}_{pN}=0;\theta\right)$, is given by

$$
f_p\left(\hat{\theta}_{p1}=0,\ldots,\hat{\theta}_{pN}=0;\theta\right)=\left\{1-\int_0^{\infty}\cdots\int_0^{\infty}\left[\text{sum of all the other terms in } f_p\left(\hat{\theta}_p;\theta\right)\right]d\hat{\theta}_{p1}\cdots d\hat{\theta}_{pN}\right\}\delta\left(\hat{\theta}_{p1}\right)\cdots\delta\left(\hat{\theta}_{pN}\right). \quad (13)
$$

Equations (11)-(13) completely define $f_p\left(\hat{\theta}_p;\theta\right)$.

3.2.2 Support-constrained estimation PDF creation

Let $\hat{\theta}_s$ be a support-constrained estimator of θ with PDF $f_s\left(\hat{\theta}_s;\theta\right)$. Unlike for the positivity constraint, the support constraint is independent of the measurement values. As a result, for a given support region, $f_s\left(\hat{\theta}_s;\theta\right)$ consists of only one term that will now be created. Let there be M elements outside of the support region, where $0 \le M < N$. We require that M be strictly less than N; otherwise, there are no elements in the support region and thus no parameters to estimate. As before, without loss of generality we assume that the first M elements of θ are the constrained elements. Following the same approach to create $f_s\left(\hat{\theta}_s;\theta\right)$ as we used to create $f_p\left(\hat{\theta}_p;\theta\right)$, we first decompose $f_s\left(\hat{\theta}_s;\theta\right)\equiv f_s\left(\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0,\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN};\theta\right)$ into a product of its conditional PDF $f_s\left(\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN}\,\middle|\,\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0;\theta\right)$ and marginal PDF $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_s\left(\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0,\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN};\theta\right)d\hat{\theta}_{s(M+1)}\cdots d\hat{\theta}_{sN}$. Mathematically, we have

$$
\begin{aligned}
f_s\left(\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0,\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN};\theta\right) &= f_s\left(\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN}\,\middle|\,\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0;\theta\right) \\
&\times\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_s\left(\hat{\theta}_{s1}=0,\ldots,\hat{\theta}_{sM}=0,\hat{\theta}_{s(M+1)},\ldots,\hat{\theta}_{sN};\theta\right)d\hat{\theta}_{s(M+1)}\cdots d\hat{\theta}_{sN}
\end{aligned}
\quad (14)
$$

For the same reasons as given in Subsection 3.2.1, we set the support-constrained conditional PDF equal to the unconstrained conditional PDF $f_u\left(\hat{\theta}_{u(M+1)},\ldots,\hat{\theta}_{uN}\,\middle|\,\hat{\theta}_{u1},\ldots,\hat{\theta}_{uM}=0;\theta\right)$, where

given by Eq.(17), the support-only case. When the support region contains all the elements of $\hat{\theta}_u$, K = 0 and $f_{ps}\left(\hat{\theta}_{ps};\theta\right)$ is given by Eqs.(11)-(13) depending upon the actual values of the elements of $\hat{\theta}_u$.

3.3 Estimation PDF extensions

The estimation PDF creation method described in Subsection 3.2 was developed for estimators using every combination of the positivity and support constraints. The essence of the positivity and support constraints is that estimated parameter values are constrained to be zero if they are either non-zero (support) or less than zero (positivity). However, the positivity constraint can be generalized to the situation where each parameter value is known to be either bounded below, bounded above, or bounded both below and above by constants that can vary for each parameter. Similarly, the support constraint can be generalized to the situation where some of the true parameter values are already known exactly. The necessary modifications are to replace the zero values in the conditional PDF expressions with the known equality values or the inequality boundary values and to replace the integration limits in the marginal PDF expressions to be consistent with the known inequality boundary values.

Equality or inequality constraints that consist of combinations of true parameter values cannot be accommodated using this estimation PDF creation method, though. To demonstrate why this is so, consider the case where the sum of the true parameter values is known to be less than or equal to a known value. There are an uncountably infinite number of combinations of estimated parameter values that satisfy this constraint. Because the method described in Subsection 3.2 creates a separate expression for each possible combination of parameter values that satisfy the known constraint, this inequality constraint obviously cannot be accounted for. For this reason, the estimation PDF creation method can be used only when constraints on individual parameter values are known.

Finally, the estimation PDF creation method can be extended to non-invertible FIMs by recasting the estimation problem to be invertible and carrying out the method on this recast problem. For example, consider the case where the parameters to be estimated correspond to a discretized time signal and the measurements are a low-pass-filtered and noisy version of the discretized time signal. The estimation problem can be recast by setting the true set of parameters to be estimated to a low-pass-filtered version of the original true set of parameters. However, the known equality and inequality constraints on the true parameter values must be able to be transformed to the low-pass-filtered parameter set for this approach to be successful. For example, a positivity constraint on the original set of parameters is still valid for a low-pass-filtered set of parameters if the low-pass filter is chosen to be non-negative. A support constraint, on the other hand, will not be strictly valid for a low-pass-filtered version of a set of parameters that have finite support. However, a broadened support region can contain virtually all of the low-pass-filtered parameter energy and is often applied in this situation.

3.4 Estimation PDF creation example

In this subsection we apply the estimation PDF creation method described in Subsection 3.2 to a specific measurement model and three different estimators in order to demonstrate how to carry out the creation process. First, the measurement model will be described. This model is chosen to be simple in order to make the concepts described in the previous subsections straightforward to illustrate. Let the measurement vector **y**, the parameter vector **θ**, and the noise vector **η** be two-dimensional and be related by the following equation

$$\mathbf{y} = \boldsymbol{\theta} + \boldsymbol{\eta}, \tag{19}$$

where the noise PDF $f_N(\boldsymbol{\eta})$ is zero mean and Gaussian with a covariance matrix given by

$$\mathbf{C} = \begin{bmatrix} \sigma_1^2 + \sigma_2^2 & \sigma_1^2 - \sigma_2^2 \\ \sigma_1^2 - \sigma_2^2 & \sigma_1^2 + \sigma_2^2 \end{bmatrix}. \tag{20}$$

From Eq.(19), it can be seen that measurement PDF $f_Y(\mathbf{y};\boldsymbol{\theta})$ is equal to $f_N(\boldsymbol{\eta} = \mathbf{y} - \boldsymbol{\theta})$, where the notation $f_N(\boldsymbol{\eta} = \mathbf{y} - \boldsymbol{\theta})$ denotes the PDF $f_N(\boldsymbol{\eta})$ with **η** replaced with **y-θ**. The relative values of σ_1^2 and σ_2^2 in Eq.(20) define the measurement statistical correlation properties. When they are equal, the two components of the measurement vector **y** are uncorrelated. When $\sigma_2^2 = 0$, the two measurement components are perfectly correlated, and when $\sigma_1^2 = 0$, the two measurement components are perfectly anti-correlated. We note that this definition of **C** is not conventional. The conventional way to define the covariance matrix of two correlated random variables, each

$$f_{ps}\left(\hat{\theta}_{ps1},\hat{\theta}_{ps2};\theta\right)=\frac{f_Y\left(y_1=\hat{\theta}_{ps1},y_2=0;\theta\right)}{\int_{-\infty}^{\infty} f_Y\left(y_1,y_2=0;\theta\right)dy_1}u^+\left(\hat{\theta}_{ps1}\right)\delta\left(\hat{\theta}_{ps2}\right)$$
$$+\left[1-\frac{\int_0^{\infty} f_Y\left(y_1=\hat{\theta}_{ps1},y_2=0;\theta\right)d\hat{\theta}_{ps1}}{\int_{-\infty}^{\infty} f_Y\left(y_1,y_2=0;\theta\right)dy_1}\right]\delta\left(\hat{\theta}_{ps1}\right)\delta\left(\hat{\theta}_{ps2}\right) \quad (23)$$

Equations (21) – (23) can be used to calculate the FIMs for the three estimators using Eq.(1). The support-constrained estimator is an unbiased estimator of θ, so the support-constrained CRBs are just the diagonal elements of the inverse of the support-constrained FIM. However, both the positivity-constrained and positivity- and support-constrained estimators are biased estimators of θ, so their bias gradient matrices must be calculated using Eq.(5) and the corresponding constrained PDFs in order to calculate the biased CRBs using Eq.(6).

4. Biased CRBs for positivity and support

The CRBs corresponding to the measurement model and the three estimators described in Subsection 3.3 (positivity-constrained, support-constrained, and both positivity- and support-constrained) are analyzed in this section as functions of the single-element signal-to-noise ratio (SNR) and the degree of correlation as determined by the ratio σ_1^2/σ_2^2. The SNR of an element is defined as the noise-free parameter value divided by the square root of the noise variance. We compare the CRBs to sample variances obtained using computer simulation results and show that the CRBs are either achieved or closely approached.

The computer simulation results presented in this section use Eq.(19) as the system model with the zero-mean additive noise term having unit variance (i.e., $\sigma_1^2+\sigma_2^2=1$) and varying values of σ_1^2 and σ_2^2 that satisfy $\sigma_1^2+\sigma_2^2=1$. The values of θ were varied in order to vary the single-element SNR values. A cost-function-based estimation scheme was used to estimate θ from $\mathbf{y}$ based on the following cost function:

$$J(\alpha)=\left(\mathbf{y}-\alpha^2\right)^T\mathbf{C}^{-1}\left(\mathbf{y}-\alpha^2\right), \quad (24)$$

where $\mathbf{C}$ is given by Eq.(20), $\hat{\theta}=\alpha^2$, and the minimization was performed over α. This parameterization of $\hat{\theta}$ enforced the positivity constraint while still permitting an unconstrained (conjugate gradient) minimization routine to be used. The unconstrained estimator corresponding to Eq.(24) is just the maximum-likelihood estimator corresponding to the measurement PDF. When both positivity and support were used as constraints, one element of θ was set to zero and the minimization was carried out over the element of α that corresponded to the non-zero element of θ. When only support was used as a constraint, the α^2 term in Eq.(13) was replaced with α and the minimization was carried out as before. Each sample variance value was calculated using 10,000 estimates of $\hat{\theta}$ generated from 10,000 statistically-independent measurements for a given set of values for σ_1^2, σ_2^2, and θ.

The results for the first set of cases to be presented were generated for θ = $[0.1, 0]^T$, $[1, 0]^T$, and $[5, 0]^T$, while varying σ_1^2/σ_2^2 from 0.01 to 100. Because the noise variance in each element is one regardless of the value of σ_1^2/σ_2^2, the values for θ are also the single-element SNRs. A positivity constraint was enforced but not a support constraint. These choices of values and constraints correspond to estimating parameter values both inside and outside a support region with the use of only a positivity constraint. The resulting biased CRBs and sample variances for $\hat{\theta}_1$ (i.e., for values inside the support region) are plotted in Figure 1, where the CRBs would be unity without the positivity constraint. Notice the good agreement between the sample variances and the biased CRBs. In addition, when σ_1^2/σ_2^2 = 1, the noise values in the two measurement elements are uncorrelated, so the reductions in the biased CRBs and sample variances corresponding to $\hat{\theta}_1$ for this value of σ_1^2/σ_2^2 come entirely from zeroing out the negative values in y_1. Thus, when the SNR of y_1 is five, the probability of obtaining a negative value for y_1 is quite small, so the biased CRB for $\hat{\theta}_1$ in this case is essentially equal to the measurement variance. As the single-

variety of measurement SNRs and noise covariances. Sample variances were generated for the same set of measurement SNRs and noise covariances and shown to match closely the theoretical CRB values.

An important outcome of this work is the explicit recognition of the role that noise correlations play in determining the effectiveness of additional knowledge such as positivity or support to reduce noise in the estimated parameters. This result provides insight into how positivity and support will function to reduce the noise levels in parameter estimates where deconvolution is part of the parameter-estimation algorithm. Positivity and support are not effective in reducing noise levels in estimated parameters when the noises are spatially delta correlated, which includes many standard noise sources such as amplifier and Poisson noise. However, the deconvolution operation imposes correlations on delta-correlated noises, leading to the possibility of noise reduction in this case. For the case of a support constraint, it has been shown that meaningful noise reduction is possible when the measurement model includes non-negligible blurring.[9]

Future work includes extending the results in this paper to delta-correlated measurement noises and deconvolution. Because the method presented in this paper is a function of the measurement noise values, this extension is straightforward in concept but may result in calculations that are intractable. Therefore, future work also includes developing variations to the method presented herein that are computationally friendly. One way to significantly reduce the computational burden is to calculate the positivity-constrained PDF terms only for elements whose SNRs are less than three. If the set of parameters to be estimated has finite support, this information should always be included because the support-constrained PDF terms are easy to calculate.

Acknowledgements

The authors wish to thank the Air Force Office of Scientific Research and the Air Force Research Laboratory for their financial support that made this research possible.

References

[1] B. Porat, *Digital Processing of Random Signals* (Prentice Hall, Englewood Cliffs, 1994).

[2] A. O. Hero, III, J. A. Fessler, and M. Usman, "Exploring estimator bias-variance tradeoffs using the uniform Cramér-Rao bound," *IEEE Trans. Signal Process.* **44**, 2026-2041 (1996).

[3] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 – theory," *IEEE Trans. Med. Imaging* **MI-1**, 81-94 (1982).

[4] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory* **26**, 1285-1301 (1990).

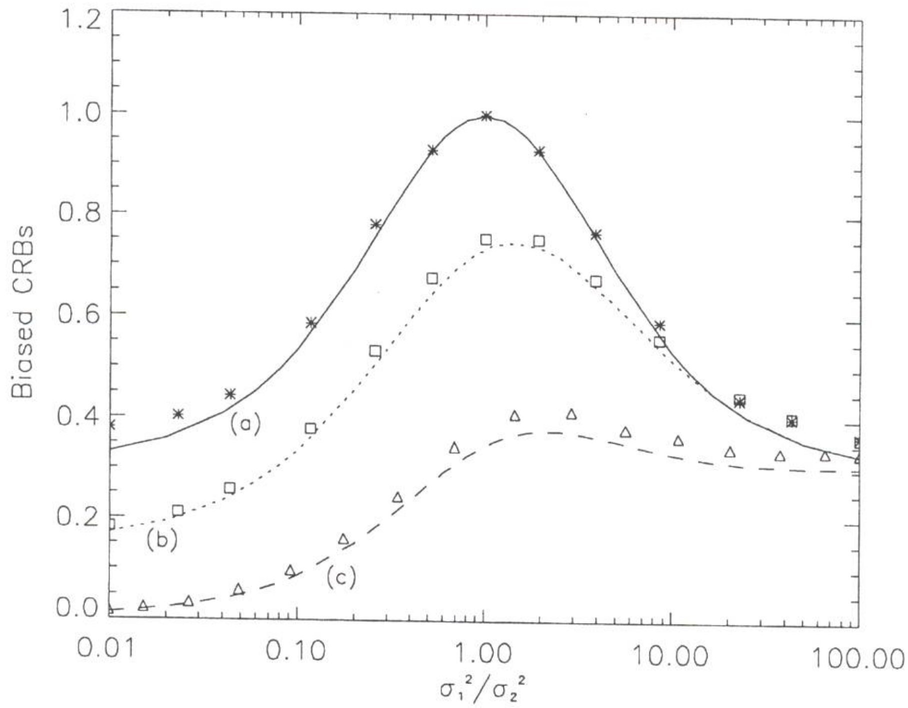
[5] D. L. Donoho, I. M. Johnstone, J. C. Hoch, and A. S. Stern, "Maximum entropy and the nearly black object," *J. R. Statist. Soc.* B **54**, 41-81 (1992)

[6] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. Signal Process.* **49**, 87-90 (2001).

[7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, 3rd ed.*, (McGraw-Hill, New York, 1991).

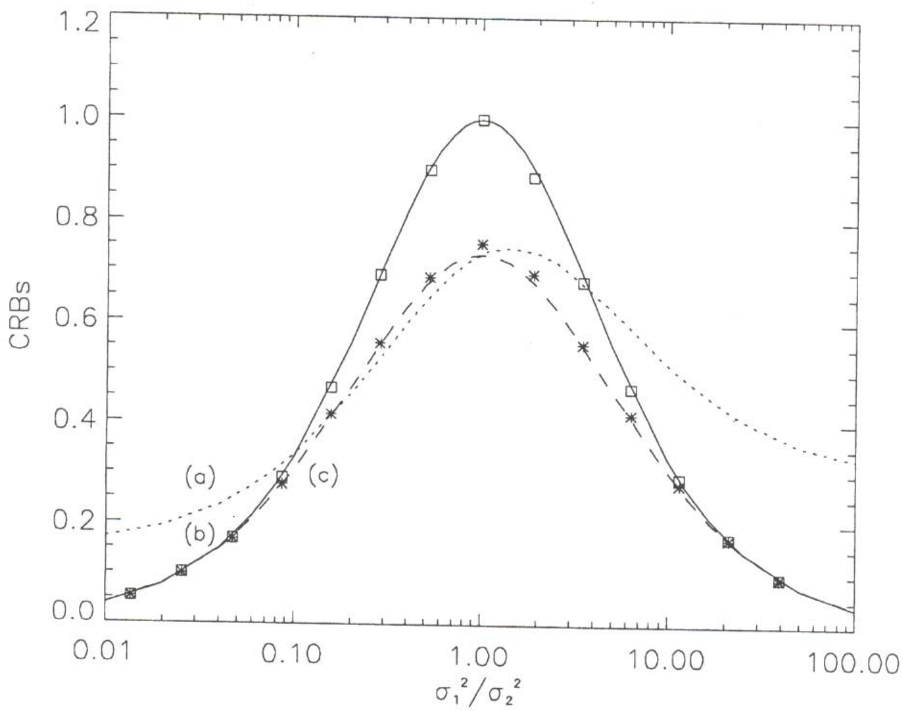
[8] Goodman, J., *Statistical Optics*, (John Wiley and Sons, New York, 1985)

[9] C. L. Matson, C. C. Beckner, and K. J. Schulze, "Fundamental limits to noise reduction in images using support - benefits from deconvolution," in *Image Reconstruction from Incomplete Data III*, P. J. Bones, M. A. Fiddy, and P. Millane, eds., *Proc. SPIE* **5562** 161-168 (2004)



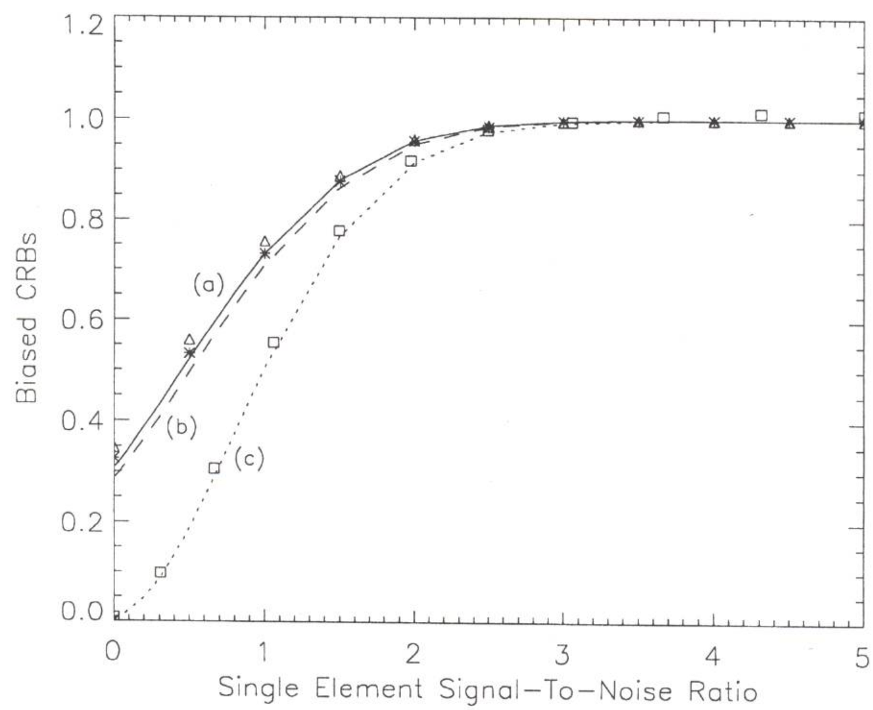
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Biased CRBs
(a)
(b)
(c)
0.01
0.10
1.00
10.00
100.00
σ_1^2/σ_2^2

Figure 1



1.2
1.0
0.8
0.6
0.4
0.2
0.0
CRBs
(a)
(b)
(c)
0.01
0.10
1.00
10.00
100.00
σ_1^2/σ_2^2

Figure 3



1.2
1.0
0.8
0.6
0.4
0.2
0.0
Biased CRBs
(a)
(b)
(c)
0
1
2
3
4
5
Single Element Signal-To-Noise Ratio

Figure 5